

STATE-OF-THE-ART PAPER

Interpreting the Results of Cost-Effectiveness Studies

David J. Cohen, MD, MSc, FACC,* Matthew R. Reynolds, MD, MSc, FACC†

Kansas City, Missouri; and Boston, Massachusetts

In developed nations, health care spending is an increasingly important economic and political issue. The discipline of cost-effectiveness (CE) analysis has developed over several decades as a tool for objectively assessing the value of new medical strategies, by simultaneously examining incremental health benefits in light of incremental costs. The underlying goal of CE research is to allow clinicians and policymakers to make more rational decisions regarding clinical care and resource allocation. This review will provide the reader with an understanding of the theoretical underpinnings of CE analysis, the types of analyses commonly performed and reported in the medical literature, some important strengths and weaknesses of different analytical approaches, and key principles in the interpretation of CE results. Key principles reviewed include the impact of analytic perspective, the importance of proper incremental comparisons, the effect of time horizon, and methods for exploring and describing uncertainty. Illustrative examples from the cardiology literature are discussed. (J Am Coll Cardiol 2008;52:2119–26) © 2008 by the American College of Cardiology Foundation

Developed nations face difficult decisions about how to allocate resources to health care, and how to prioritize spending within their health care systems. In the U.S., for example, over at least several decades, the growth in spending on health care has consistently outpaced the growth of the overall economy. While most observers agree that this trend is unsustainable long term, and is already producing political and economic problems, enacting measures to reduce the growth in health care spending has proved difficult (1).

It has long been recognized that new medical products and technologies are one important driver of increased health care costs (2–4). This realization has increasingly highlighted the need to assess the value of new clinical strategies as they are introduced, that is, to measure the benefits of tests, drugs, procedures, and medical devices relative to their costs. The discipline of cost-effectiveness (CE) analysis aims to evaluate such questions in order to inform medical decision making and health care policy.

Fundamentals of Health Economic Assessment

Health economic studies can take many forms and report a variety of possible outcomes. Typically, 1 or more new

strategies are compared against an existing standard of care with regard to the dual outcomes of clinical effectiveness and cost. One can readily visualize the possible results of such joint comparisons in a 2-dimensional plot (5)—often referred to as the “cost-effectiveness plane” (Fig. 1)—in which the standard of care occupies the origin of the graph. The new intervention(s) under study will locate themselves to the right or left of the origin if they are more or less effective than the current standard of care, and above or below the origin if they are more or less costly.

When a new intervention is both clinically superior and cost saving, it is referred to as an economically “dominant” strategy. The opposite is a “dominated” strategy. Few novel technologies will fall into either of these categories, however; the most common scenario is that a new strategy improves clinical results at increased cost. In these cases, the estimation of value is based upon calculation of a CE ratio (see the following text).

These different potential outcomes give rise to a variety of terms for individual types of health economic studies. A study aimed at establishing the least costly among clinically equivalent strategies is called a cost-minimization study, but cost-minimization studies rely on the premise that clinical equivalence has been proven, which can be difficult and sometimes controversial (6). Therefore, readers should be careful to evaluate the clinical evidence for equivalence or therapeutic interchangeability before placing much weight on the results of a cost-minimization study. CE studies, in contrast, calculate incremental costs in units of currency, while expressing clinical benefits in nonmonetary terms such as life-years gained or adverse events avoided. Cost-utility

From *Saint Luke's Mid America Heart Institute, Kansas City, Missouri; and †Beth Israel Deaconess Medical Center, Harvard Clinical Research Institute, Boston, Massachusetts. Dr. Cohen reports grant support from Eli Lilly, Cordis, Boston Scientific, Bristol-Myers Squibb/Sanofi, The Medicines Company, and Edwards Lifesciences, as well as consulting fees from Medtronic. Dr. Reynolds is supported by grant K23 HL077171 from the National Heart, Lung, and Blood Institute and also reports grant support from Edwards Lifesciences and consulting fees from Biosense Webster, and Sanofi-Aventis.

Manuscript received July 22, 2008; revised manuscript received September 2, 2008, accepted September 29, 2008.

Abbreviations and Acronyms

CE = cost-effectiveness

CRT-D = cardiac resynchronization therapy defibrillator

CRT-P = cardiac resynchronization therapy pacemaker

DES = drug-eluting stent(s)

ICER = incremental cost-effectiveness ratio

QALY = quality-adjusted life year

analyses, a subset of CE analyses, estimate effectiveness using measures that reflect individual or societal preferences for differing health states, such as quality-adjusted life years (QALYs).

As mentioned, the majority of new health strategies improve clinical results at increased cost. The generic formula for calculating a CE ratio in these cases is as follows:

$$\text{CE ratio} = \frac{\text{Cost}_{\text{New}} - \text{Cost}_{\text{Reference}}}{E_{\text{New}} - E_{\text{Reference}}}$$

where E is the effectiveness measure. Like any statistical measure, the point estimate for a CE ratio obtained using the above formula is surrounded by some degree of uncertainty, and that uncertainty may overlie more than 1 quadrant of the CE plane. Specialized methods have been developed to measure and display this uncertainty (see the following text).

No single threshold exists for deciding whether or not a CE ratio is acceptable. Obviously, a variety of considerations including the prosperity of a nation or health system would dictate what type of thresholds might be affordable. Within the U.S., where no policy-making emanates directly from CE analysis (at least at present), CE ratios of <\$50,000 per life year gained are generally considered attractive, and >\$100,000 per life year gained are generally consid-

ered unattractive, but these are rough guidelines at best and have been criticized as outdated and artificially low (7,8).

In other nations such as the United Kingdom and Australia, however, health economic studies are an integral component of the evaluation of any new medical treatment, and explicit CE thresholds (e.g., £30,000 per QALY gained) have been promulgated, though not always as absolute standards (9). Health economic studies, therefore, must be interpreted within the appropriate geopolitical context, and CE ratios, when published, are often compared with those from previous studies of other interventions that were accepted (or not) at clinical and policy levels. While this type of “relativism” has flaws (e.g., many accepted practices have never been subjected to careful health economic scrutiny), it does provide a quantitative and objective perspective on the value of new technologies and treatment strategies.

Several corollaries can be discerned from the formula for calculating the CE ratios. First, it should be self-evident that treatments that increase net cost compared with the available alternatives can only be cost-effective if they provide a net clinical benefit. Second, cost-saving strategies tend to be cost-effective only if they are at least close to clinically neutral, but this depends on the CE threshold. Thus, cost-saving and -effective are not synonymous terms, and it is possible for a less expensive and slightly less effective strategy to be preferred on health economic grounds, particularly in settings where resources are highly constrained (10). In contrast, even interventions that are quite expensive may be reasonably cost-effective if they result in significant gains in life expectancy and the CE threshold is high.

Measuring costs. In 1996, the U.S. Panel on Cost-Effectiveness in Health and Medicine codified the preferred assignment of costs and benefits to the numerator and denominator of CE ratios in an effort to foster methodological consistency across studies (11). Important categories of cost that should be measured in health economic studies include the direct medical costs associated with each clinical strategy; “induced” or downstream costs incurred (e.g., those associated with late complications) or avoided (e.g., subsequent hospital admissions) due to the strategy; and certain indirect costs, such as time and travel for family members who often act as unpaid caregivers. Well-conducted analyses must fairly and accurately account for each of these costs. In the decentralized U.S. health care system, this often requires the careful collection and review of claims data, the fastidious collection of resource utilization data (which can be converted to costs using representative price weights for each item), or both.

The U.S. panel further recommends that the economic impact of illness on individual patients (e.g., lost wages from disability or death) be incorporated in the denominator of a CE ratio (as reflected in life expectancy or quality-adjusted life expectancy) and not in the numerator. Such productivity costs, while sometimes important for understanding the full

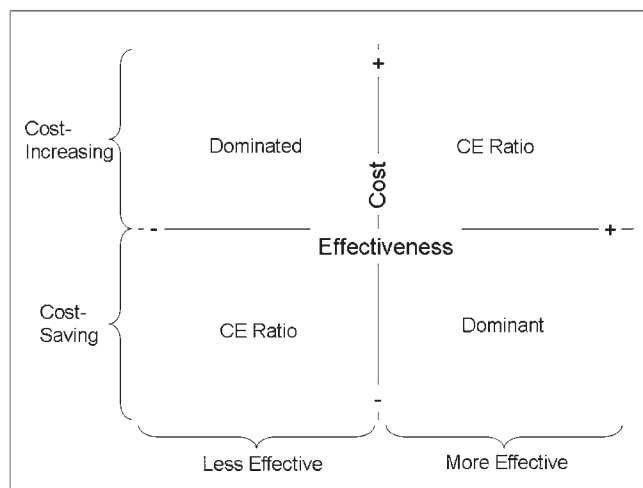


Figure 1 The CE Plane

With some reference strategy occupying the origin of the graph, a cost-effectiveness (CE) study can plot the incremental costs (y-axis) and benefits (x-axis) of alternative strategies, relative to this reference, in 2-dimensional space. The area above the horizontal is cost-increasing, and to the right of the vertical, clinically beneficial. When a new strategy adds both benefits and costs (upper right-hand quadrant) or reduces both (lower left-hand quadrant), a CE ratio must be calculated to judge benefits relative to costs.

economic impact of an illness, are, therefore, frequently not included in contemporary CE studies.

Measuring effectiveness. There are, theoretically, few constraints on what measure of effectiveness is used in the denominator of a CE ratio, although some measures clearly have more appeal than others. Changes in life expectancy generally trump other outcomes and form the focus of many health economic studies in cardiology (e.g., for implantable defibrillators or coronary revascularization). Advantages of this approach include the unquestioned value that patients attach to improved survival and the fact that mortality rates are readily measured in many clinical trials. Investigators have also used the avoidance of adverse events, such as ischemia-driven repeat revascularization procedures (12), as effectiveness measures in CE studies; this approach appears most acceptable when the adverse events are associated with measurable decrements in quality of life.

Some desirable interventions may not alter life expectancy but still offer value through reduction or avoidance of symptoms and improvement in quality of life, and others may significantly alter both the quantity and quality of life. It is here that cost-utility analyses are recommended, with QALYs serving as the preferred measure of effectiveness (11). Authorities favor the use of QALYs in CE studies because, at least in theory, they can be measured across a wide variety of health conditions. To calculate QALYs, one must measure utility weights, which reflect an individual's preference for a given health state on a scale ranging from 1.0 (perfect health) to 0 (death) (13). A person's (or population's average) utility may change over time and through the course of an illness. QALYs are calculated as utility multiplied by the length of time (in years) spent in the health state corresponding with that utility, summed over time (Fig. 2).

The chief drawback to using QALYs for CE analysis lies in the methods available for measuring utility. Gold-standard methods of directly eliciting utilities from patients are strongly grounded in economic theory but difficult and time-consuming to apply in practice (14,15). For this reason, investigators more often use indirect methods, in which study participants complete generic health state classification surveys (e.g., the EuroQol [16] or the Medical Outcomes Study Short-Form 36 [17]) that, in turn, have previously estimated utilities from reference populations for each health state defined by the survey (18,19). Due to the intricacies of utility assessment and conflicting guidance on the topic, CE studies vary widely in their approach to quality-of-life adjustment (20), and all too often the data needed for proper quality adjustment simply are not available. We believe that the widespread availability of validated, multilingual instruments for assessment of population utilities is an important recent advance that should lead to increasing consistency and validity in health economic studies.

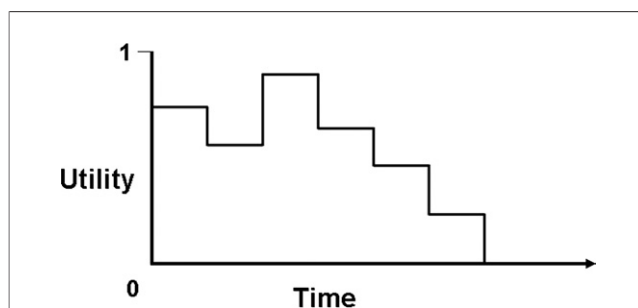


Figure 2 Quality-Adjusted Life Expectancy

To calculate quality-adjusted life expectancy, the time spent in a particular health state (typically measured in years) is multiplied by the utility weight (possible range of 0 to 1) for that health state, and these products are then summed over time. Quality-adjusted life expectancy is thus represented by the area under the hypothetical “curve” in this figure.

Types of Health Economic Studies

Readers will encounter 3 basic kinds of health economic studies, each with its own distinct strengths and limitations (Table 1). Trial-based studies (21,22) generally benefit from careful and accurate data collection; from randomization, which minimizes bias and confounding; and from the rigorous adjudication of end points. However, important aspects of clinical trials may differ from the “real world” in terms of patient selection and recruitment, clinical management, or other factors that are important to economic outcomes. Perhaps more importantly, clinical trials are often limited by finite (and potentially short) time horizons and unequal follow-up duration within groups. If the trial duration is not sufficiently long to capture all of the pertinent clinical and economic ramifications of the strategies under study, then the estimation of CE may be biased (see the following text). Finally, pure trial-based analyses tend not to incorporate data from external sources, exposing the results to potentially greater uncertainty than if evidence from other trials was considered.

Some economic studies derive entirely from disease-simulation models (23). Common approaches in the medical literature include Markov models (24) and discrete event simulation (25). Models are mathematical structures that represent the key aspects of the strategies under study, and can incorporate data from a wide variety of sources as inputs. Models can estimate likely CE outcomes when clinical trials are not feasible, or not yet complete. In addition, model-based analyses can incorporate multiple competing strategies, which are generally impractical to examine in a clinical trial setting. However, models generally require simplifying assumptions, and ultimately reflect the accuracy of the source data on which they are built—good or poor. Finally, models can incorporate the results of systematic overviews of therapeutic efficacy (i.e., meta-analyses), thus overcoming limitations introduced by over-reliance on the results of any single trial. When conducted well, modeling studies make their assumptions transparent,

Table 1 Types of Health Care Economic Studies

| | Definition | Advantages | Disadvantages | Examples From ICD Literature |
|-------------|--|--|--|-------------------------------|
| Trial-based | Economic analysis performed alongside clinical trial | <ul style="list-style-type: none"> Precise measurement of costs and benefits during trial and within the same population Randomization minimizes bias | <ul style="list-style-type: none"> Trial conditions may differ from “real life” Limited time horizon | MADIT (21) AVID (22) |
| Model-based | Analytic structure developed using best available evidence | <ul style="list-style-type: none"> Predict likely outcomes when trials not feasible Can incorporate inputs from multiple sources Can consider strategies that have not been directly tested in head-to-head trials Identify key areas of uncertainty | <ul style="list-style-type: none"> Require simplifying assumptions Key model inputs may be flawed or unknown | Sanders et al. (23) |
| “Hybrid” | “In-trial” results extrapolated using modeling techniques | <ul style="list-style-type: none"> Trial data informs model Adjust time horizon for more “fair” comparison | <ul style="list-style-type: none"> Modeling portion has same limitations as pure model-based study | MADIT II (8) SCD-HeFT (26) |

AVID = Antiarrhythmics Versus Implantable Defibrillators; ICD = implantable cardioverter-defibrillator; MADIT = Multicenter Automatic Defibrillator Implantation Trial; SCD-HeFT = Sudden Cardiac Death in Heart Failure Trial.

test the impact of key assumptions, and, in so doing, may identify key areas of uncertainty on which future research should focus.

Increasingly, economic studies incorporating elements of both trial- and model-based methodologies have been reported (8,26). These hybrid studies can address the limitations of trial-based analysis—in particular, the issue of truncated follow-up—by extending the results of the study through time, generating a range of plausible projections of longer-term outcomes. While those projections are potentially subject to some of the same criticisms as purely model-based studies, the hybrid approach can take advantage of the carefully collected in-trial data to inform the modeling effort.

Key Principles in the Interpretation of CE Studies

Analytic perspective. One of the most important considerations in interpreting CE research is the analytic perspective of the study. Most health systems are structured such that multiple parties are involved in the delivery, payment, and receipt of care. Each stakeholder (or group of stakeholders), following their own incentives (e.g., to maximize health, maximize revenue, or minimize expenditure), may, thus, have very different views on what represents optimal policy for a particular intervention.

Table 2 illustrates the importance of perspective by considering differing possible views on the usage of drug-

eluting stents (DES) for patients with coronary artery disease undergoing percutaneous coronary intervention. In this hypothetical example, based on the initial reimbursement policy for DES after their approval, the potentially disparate incentives, obligations, and constraints of the various parties would lead—at least in theory—to different preferences for one strategy versus another.

The standard recommendation for CE studies in medicine is to use the most inclusive perspective possible, so as to incorporate the potential benefits, harms, and costs for all parties involved. This defines the societal (or health system) perspective, which flows from the desire for CE studies to inform policy making at the broadest levels. Seen in this light, CE analyses are less concerned with individual winners and losers of a particular strategy (e.g., surgeons vs. cardiologists or hospitals vs. insurers), but rather with the more expansive aim of understanding the global balance between societal costs and societal benefits. Some have argued, however, that this approach is incomplete, and that a fully transparent accounting of CE should demonstrate explicitly the effect on each of the individual stakeholders. This is likely one important reason that traditional CE analyses taking the societal perspective have not been more widely used in policymaking.

Incremental comparisons. CE ratios are often reported as “*incremental* cost effectiveness ratios” (“iCERs” or “ICERs”),

Table 2 Theoretical Impact of Alternative Perspectives on Cost-Effectiveness of DES

| Perspective | Objective | Preference for DES |
|----------------------------|---|---|
| Patient | Maximize personal health benefits | <ul style="list-style-type: none"> DES for every lesion (assuming the patient has health insurance and out-of-pocket expenses are minimal) |
| Hospital | Maximize margin (revenue minus costs) for each episode of care | <ul style="list-style-type: none"> DES > BMS for discrete lesions that can be treated with 1 stent (incremental reimbursement covers cost of ~1.5 DES) Avoid DES for multivessel disease (especially vs. CABG) |
| CMS/insurer | Maintain health of beneficiaries with smallest possible growth in cost | <ul style="list-style-type: none"> DES reasonable for patients currently receiving CABG |
| Society/health care system | Maximize societal health benefits within constraint of overall spending limit | <ul style="list-style-type: none"> DES for patients where the expected clinical benefits are worth the additional cost |

BMS = bare-metal stent(s); CABG = coronary artery bypass grafting; CMS = Centers for Medicare & Medicaid Services; DES = drug-eluting stent(s).

with the “i” emphasizing the notion that CE is not an inherent property of any one medical technology. Rather, CE can only be estimated by the direct comparison of one clinical strategy with another. An important tenet in the calculation of “iCERs,” dictated by the economic theory underlying health economics research, is that each relevant strategy should be compared with the next best alternative, based on the economic concept of “opportunity costs” (11).

Failure to make incremental comparisons with each relevant strategy can lead to distortions in the calculation of CE ratios and potentially erroneous conclusions. This is exemplified by the CE analysis of the COMPANION (Comparison of Medical Therapy, Pacing, and Defibrillation in Heart Failure) trial (27)—a 3-armed randomized trial that compared medical therapy with cardiac resynchronization pacemakers (CRT-Ps) or defibrillators (CRT-Ds) in heart failure patients. As shown in Figure 3A, the analysts made separate comparisons of the CE of CRT-Ps and CRT-Ds with the “optimal medical therapy” control group. Both of these CE ratios appeared favorable. An editorialist, however, pointed out the omission of a comparison the trial was not designed specifically to address, but that was nonetheless of interest: CRT-Ps versus CRT-Ds (28). The significantly greater iCER for CRT-Ds when compared with CRT-Ps (Fig. 3B) raises important questions about the incremental value of the more expensive technology, and suggests that, under certain budgetary conditions (i.e., below certain CE thresholds), the modestly less effective strategy of CRT-P might actually be preferred.

Time horizon. CE studies can be exquisitely sensitive to the time horizon of analysis. Ideally, the time horizon of a CE study should cover the entire period over which the interventions may have an effect on either clinical or economic outcomes. As noted previously, this is a potential weakness of purely trial-based analyses, particularly if a strategy under study involves primarily up-front expenditure, but provides clinical benefits that extend beyond the duration of the trial—a common scenario for many preventive strategies. In such cases, the incremental cost comparisons for the trial may be roughly accurate, but the cumulative incremental benefits may be significantly underestimated (because much of the benefit occurs beyond the time frame observed during the trial), resulting in artificially high CE ratios.

The CE studies from 2 recent implantable cardioverter-defibrillator trials demonstrate these concepts. For both the MADIT (Multicenter Automatic Defibrillator Implantation Trial) II and SCD-HeFT (Sudden Cardiac Death in Heart Failure Trial) studies, the up-front expenditures of device implantation coupled with the moderate length (3 to 5 years) of the trials translated into fairly high CE ratios using empirical in-trial data (\$127,000 to \$235,000 per life-year gained). To address this issue, both groups of investigators also calculated CE ratios based on longer-term projections of survival and costs of their study cohorts, and found that the resulting CE ratios decreased to ~\$60,000 to

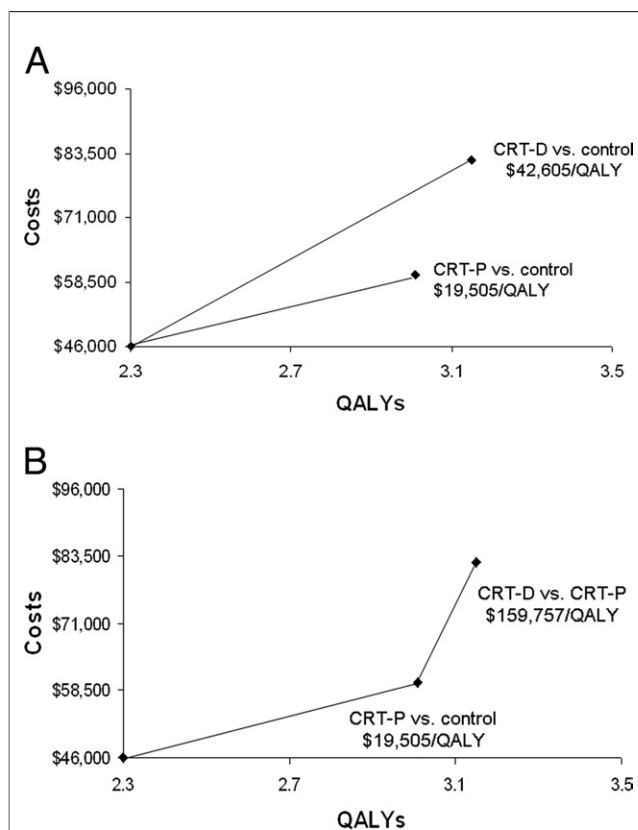


Figure 3 Absolute Versus Incremental CE

The incremental costs (y-axis) and effectiveness (x-axis, in quality-adjusted life years [QALYs]) for the 2 experimental arms of the COMPANION (Comparison of Medical Therapy, Pacing, and Defibrillation in Heart Failure) trial are plotted versus the control group of optimal medical therapy (27). In **A**, the reported separate cost-effectiveness (CE) ratios for cardiac resynchronization therapy pacemakers (CRT-Ps) and cardiac resynchronization therapy defibrillators (CRT-Ds) versus the control group both appear to be attractive from a U.S. perspective. The CE theory, however, dictates that each alternative is compared to the next best. When this is done, as shown in **B** (28), the incremental CE ratio for CRT-D versus CRT-P appears much larger.

\$80,000 per life year gained at 12 years (8,26), and ~\$40,000 per life-year gained in a lifetime model (26).

It is also possible for studies with limited time horizons to underestimate CE ratios, leading to an overly optimistic view of CE. This outcome might occur if a therapy requires continuing long-term expense with diminishing clinical returns over time. For example, analysis from the CURE (Clopidogrel in Unstable angina to prevent Recurrent Events) trial (29) found that the addition of clopidogrel to daily aspirin for up to 1 year after an acute coronary syndrome was highly cost-effective, with a CE ratio of <\$10,000 per life year. In contrast, a separate modeling study (30) explored the implications of longer-term therapy for the same indication, and found that by 3 to 5 years, continued clopidogrel treatment resulted in highly unfavorable CE ratios because incremental benefit changed little over time, while incremental costs increased substantially, largely due to the continued cost of the drug itself.

There is no single time horizon applicable to all CE studies. We believe that, for interventions that affect mortality, the most appropriate time frame for analysis should be the patient's lifetime. While a lifetime perspective creates analytic challenges for investigators, including the potential need for highly uncertain extrapolations, it ensures that all important long-term costs and benefits are considered. For interventions where all (or at least most) expenditures and benefits occur in the near-term, fairly short time horizons may be appropriate. One recent example of such an analysis is the case of DES versus bare-metal stents for patients undergoing percutaneous coronary intervention. In this case, both the benefits and incremental costs of DES largely accrue during the first year of follow-up (when restenosis generally occurs) and a 1-year, trial-based time horizon was reasonable (12,31). More recently, however, studies suggesting increased very late stent thrombosis with DES have raised questions about the validity of such a short-term analytic perspective (32).

Uncertainty. All empirical comparisons carry some amount of uncertainty. In clinical studies, we generally describe this uncertainty with familiar measures such as confidence intervals, *p* values, and power. Unique features of economic data and CE studies require additional methods for measuring and expressing uncertainty.

Particularly in the context of modeling, CE studies may include many individual parameters that are poorly defined or even completely unknown, thus requiring the analyst to make explicit assumptions about their values. The impact of assumed or uncertain individual parameters on the overall results of health economic studies must be systematically evaluated—a process known as uncertainty or sensitivity analysis (33). In a sensitivity analysis, model results are recalculated as important model parameters are varied across a plausible range of values. Sensitivity analyses not only point out which parameters do or do not significantly influence overall results, but can also be used to estimate threshold values above or below which one strategy becomes preferred over another.

The greater the number of uncertain parameters in a study, the more cumbersome sensitivity analysis becomes to conduct and report. Moreover, sensitivity analyses of single parameters also fail to communicate the overall uncertainty of a modeled result. To address these problems, sophisticated methods, such as probabilistic sensitivity analysis (also known as second-order Monte Carlo simulation), have been developed that allow investigators to simultaneously vary any number of model inputs at once and thereby assess the true impact of the joint uncertainty in each parameter on a model's overall findings (34,35). These techniques help to establish the confidence in a model's conclusions by reporting the proportion of iterations that favor one strategy over another.

For trial-based analyses, specialized methods are also required to express the uncertainty around point estimates for CE ratios, since neither the calculation nor interpretation of confidence intervals for CE ratios are straightforward.

Bootstrap resampling has emerged as one particularly useful technique for handling this type of uncertainty (36,37). The bootstrap method involves creating a “dummy” dataset by resampling with replacement (i.e., randomly selecting 1 patient at a time) from the original dataset and repeating this random patient selection until the dummy dataset reaches the same size as the original. The CE ratio is then recalculated from the dummy dataset, and the entire process is repeated many (e.g., 1,000) times. The average CE ratio, over many bootstrap iterations, should approximate the point estimate from the trial data, but when the result of each iteration is plotted on the CE plane, the results appear as a “cloud” of possible outcomes (Fig. 4), reflecting the variability within the original study sample.

Once the bootstrap resampling calculations are completed, the distribution of the various points in the “cloud” can be analyzed in several instructive ways. First, confidence intervals for incremental costs, incremental effectiveness, and the joint distribution of the 2 can be generated. Furthermore, the proportion of points falling in the different quadrants of the CE plane can be measured. Finally, the proportion of incremental CE ratios falling above or below any hypothetical threshold can be reported.

Since the “optimal” threshold for CE ratios has never been agreed upon, and would vary from place to place anyway, a currently favored approach to visualize the information obtained from bootstrap resampling of study results is the construction of CE acceptability curves (38). In these graphs (Fig. 5) (39), the probability that the intervention under investigation would be economically acceptable given a specific CE threshold (i.e., societal willingness to pay) is plotted on the y-axis over a wide range of possible thresholds, spread along the x-axis. CE acceptability curves provide readers with a rapid and understandable summary of the uncertainty in a study's CE point estimate, the thresholds where 1 strategy becomes favored over others, and the confidence that specific thresholds of interest have or have not been met.

Limitations of CE Research

CE research is meant to be a source of unbiased information for medical decision making and policy setting, for use in broad applications such as the development of clinical guidelines or reimbursement policy. At their best, CE studies provide insight into the tradeoffs and consequences of certain choices that would not be apparent through assessment of clinical outcomes alone. In general, however, the information obtained from CE studies is not well suited to clinical decision making at the individual patient level. Nor is CE data sufficient, by itself, for making complex resource allocation decisions, as health economic studies cannot on their own incorporate all of the values—such as equity, feasibility, or overall budgetary impact—that may be important. Total budget impact tends to be particularly important for technologies where the absolute cost of adoption—whether due to a high per-patient

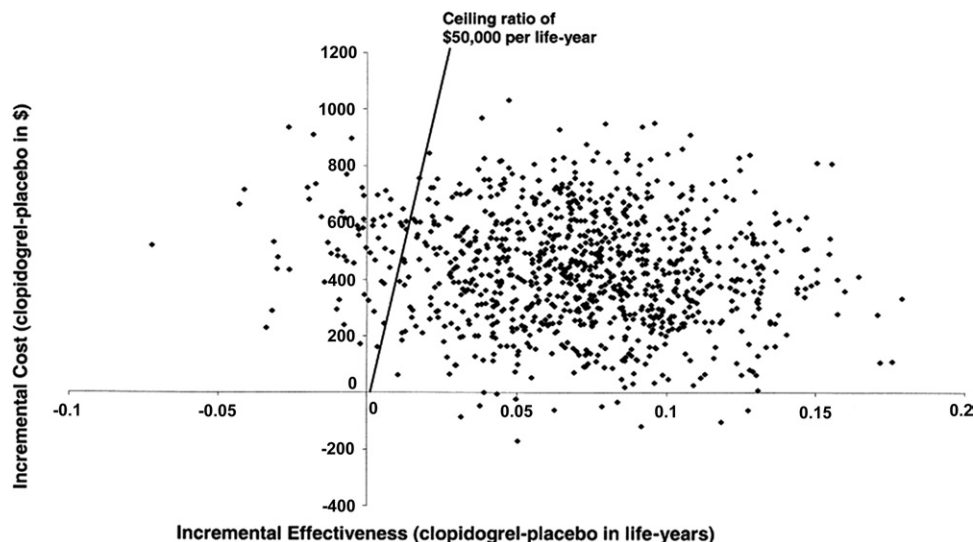


Figure 4 Results of Bootstrap Resampling in a Trial-Based CE Study

The joint distribution of projected lifetime differences in costs (y-axis) and life-expectancy (x-axis) based on the CURE (Clopidogrel in Unstable angina to prevent Recurrent Events) trial population were recalculated over 5,000 replications of the study data using the bootstrap resampling method and plotted on the cost-effectiveness (CE) plane. **Each point in the scatterplot** represents 1 bootstrap iteration. Data used, with permission, from Weintraub et al. (29).

implementation cost or due to a large number of affected individuals—is substantial.

Additional barriers have prevented the more explicit use of CE data in the development of coverage and reimbursement policy. These include political obstacles, for example, the U.S. Medicare program has no statutory mandate to

examine CE and has resisted attempts to change this (40). In addition, there are often valid concerns about the accuracy and transparency of CE data (41), and even the best studies remain subject to limitations. Finally, it is increasingly apparent that universal adoption of all new medical technologies deemed “cost-effective” by conventional criteria may have problematic budgetary consequences for important stakeholders or for health systems in general. For this reason, regulators outside the U.S. are increasingly requiring budget impact analyses along with CE studies when assessing new therapies (42). Though the processes differ, it is also clear that national coverage decisions undertaken by Medicare involve more careful scrutiny of clinical effectiveness when the financial stakes of the decision are large.

Despite these limitations we believe CE analysis will continue to grow in importance. As scientific and clinical laboratories develop new technologies to benefit our patients, both the need for investigators capable of conducting economic assessments and the need for clinicians and policymakers to understand and critically appraise CE literature will grow as well.

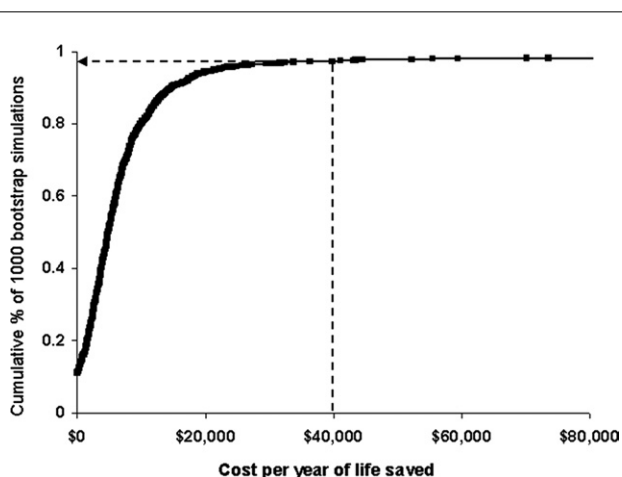


Figure 5 CE Acceptability Curve

In this example, the likelihood that use of an embolic protection device during percutaneous coronary intervention of a vein graft (vs. no distal protection) is cost-effective is shown graphically across a range of theoretical cost-effectiveness (CE) thresholds. For each CE threshold (increasing from **left to right**) on the x-axis, the proportion of bootstrap iterations having a CE ratio at or below that threshold is plotted on the y-axis. As indicated by the **arrow**, 97.3% of the bootstrapped CE ratios were <\$40,000 per year of life gained. Reprinted, with permission, from Cohen et al. (39).

Reprint requests and correspondence: Dr. David. J. Cohen, Saint Luke’s Mid America Heart Institute, 4401 Wornall Road, Kansas City, Missouri 64111. E-mail: dcohen@saint-lukes.org.

REFERENCES

1. Mongan JJ, Ferris TG, Lee TH. Options for slowing the growth of health care costs. *N Engl J Med* 2008;358:1509–14.

2. Bodenheimer T. High and rising health care costs. Part 2: technologic innovation. *Ann Intern Med* 2005;142:932–7.
3. Cutler DM, Rosen AB, Vijan S. The value of medical spending in the United States, 1960–2000. *N Engl J Med* 2006;355:920–7.
4. Goldman DP, Shang B, Bhattacharya J, et al. Consequences of health trends and medical innovation for the future elderly. *Health Affairs* 2005;24:R5–17.
5. Mark DB. Visualizing cost-effectiveness analysis. *JAMA* 2002;287:2428–9.
6. Ware JH, Antman EM. Equivalence trials. *N Engl J Med* 1997;337:1159–61.
7. Ubel PA, Hirth RA, Chernew ME, Fendrick AM. What is the price of life and why doesn't it increase at the rate of inflation? *Arch Intern Med* 2003;163:1637–41.
8. Zwanziger J, Hall WJ, Dick AW, et al. The cost effectiveness of implantable cardioverter-defibrillators: results from the Multicenter Automatic Defibrillator Implantation Trial (MADIT)-II. *J Am Coll Cardiol* 2006;47:2310–8.
9. Williams I, McIver S, Moore D, Bryan S. The use of economic evaluations in NHS decision-making: a review and empirical investigation. *Health Technol Assess* 2008;12:1–196.
10. Goldie SJ, Gaffkin L, Goldhaber-Fiebert JD, et al. Cost-effectiveness of cervical-cancer screening in five developing countries. *N Engl J Med* 2005;353:2158–68.
11. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-Effectiveness in Health and Medicine*. New York, NY: Oxford University Press, 1996.
12. Cohen DJ, Bakhai A, Shi C, et al. Cost-effectiveness of sirolimus-eluting stents for treatment of complex coronary stenoses: results from the Sirolimus-Eluting Balloon Expandable Stent in the Treatment of Patients With De Novo Native Coronary Artery Lesions (SIRIUS) trial. *Circulation* 2004;110:508–14.
13. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med* 1977;296:716–21.
14. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986;5:1–30.
15. Mark DB, Hlatky MA. Medical economics and the assessment of value in cardiovascular medicine: part I. *Circulation* 2002;106:516–20.
16. The EuroQol Group. Euro-Qol: a new facility for measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
17. Ware JE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
18. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol group. *Ann Med* 2001;33:337–43.
19. Brazier J, Roberts J, Deverill M. The estimation of a preference based measure of health from the SF-36. *J Health Econ* 2002;21:271–92.
20. Brauer CA, Rosen AB, Greenberg D, Neumann PJ. Trends in the measurement of health utilities in published cost-utility analyses. *Value Health* 2006;9:213–8.
21. Mushlin AI, Hall WJ, Zwanziger J, et al. The cost-effectiveness of automatic implantable cardiac defibrillators: results from MADIT. Multicenter Automatic Defibrillator Implantation Trial. *Circulation* 1998;97:2129–35.
22. Larsen G, Hallstrom A, McAnulty J, et al. Cost-effectiveness of the implantable cardioverter-defibrillator versus antiarrhythmic drugs in survivors of serious ventricular tachyarrhythmias: results of the Antiarrhythmics Versus Implantable Defibrillators (AVID) economic analysis substudy. *Circulation* 2002;105:2049–57.
23. Sanders GD, Hlatky MA, Every NR, et al. Potential cost-effectiveness of prophylactic use of the implantable cardioverter defibrillator or amiodarone after myocardial infarction. *Ann Intern Med* 2001;135:870–83.
24. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993;13:322–8.
25. Caro JJ. Pharmacoeconomic analyses using discrete event simulation. *Pharmacoeconomics* 2005;23:323–32.
26. Mark DB, Nelson CL, Anstrom KJ, et al. Cost-effectiveness of defibrillator therapy or amiodarone in chronic stable heart failure. Results from the Sudden Cardiac Death in Heart Failure Trial (SCD-HeFT). *Circulation* 2006;114:135–42.
27. Feldman AM, de Lissovoy G, Bristow MR, et al. Cost effectiveness of cardiac resynchronization therapy in the Comparison of Medical Therapy, Pacing and Defibrillation in Heart Failure (COMPANION) trial. *J Am Coll Cardiol* 2005;46:2311–21.
28. Hlatky MA. Cost effectiveness of cardiac resynchronization therapy. *J Am Coll Cardiol* 2005;46:2322–4.
29. Weintraub WS, Mahoney EM, Lamy A, et al. Long-term cost-effectiveness of clopidogrel given for up to one year in patients with acute coronary syndromes without ST-segment elevation. *J Am Coll Cardiol* 2005;45:838–45.
30. Schleinitz MD, Heidenreich PA. A cost-effectiveness analysis of combination antiplatelet therapy for high-risk acute coronary syndromes: clopidogrel plus aspirin versus aspirin alone. *Ann Intern Med* 2005;142:251–9.
31. Bakhai A, Stone GW, Mahoney E, et al. Cost effectiveness of paclitaxel-eluting stents for patients undergoing percutaneous coronary revascularization: results from the TAXUS-IV trial. *J Am Coll Cardiol* 2006;48:253–61.
32. Garg P, Cohen DJ, Gaziano T, Mauri L. Balancing the risks of restenosis and stent thrombosis in bare-metal versus drug-eluting stents: results of a decision analytic model. *J Am Coll Cardiol* 2008;51:1844–53.
33. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Econ* 1994;3:95–104.
34. Critchfield GC, Willard KE. Probabilistic analysis of decision trees using Monte Carlo simulation. *Med Decis Making* 1986;6:85–92.
35. Halpern EF, Weinstein MC, Hunink MG, Gazelle GS. Representing both first- and second-order uncertainties by Monte Carlo simulation for groups of patients. *Med Decis Making* 2000;20:314–22.
36. Efron B. *An Introduction to the Bootstrap*. New York, NY: Chapman & Hall, 1993.
37. Hunink MG, Bult JR, de Vries J, Weinstein MC. Uncertainty in decision models analyzing cost-effectiveness: the joint distribution of incremental costs and effectiveness evaluated with a nonparametric bootstrap method. *Med Decis Making* 1998;18:337–46.
38. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects, and C/E-ratios alongside a clinical trial. *Health Econ* 1994;3:309–19.
39. Cohen DJ, Murphy SA, Baim DS, et al. Cost-effectiveness of distal embolic protection for patients undergoing percutaneous intervention of saphenous vein bypass grafts: results from the SAFER trial. *J Am Coll Cardiol* 2004;44:1801–8.
40. Neumann PJ, Rosen AB, Weinstein MC. Medicare and cost-effectiveness analysis. *N Engl J Med* 2005;353:1516–22.
41. Tunis SR. Why Medicare has not established criteria for coverage decisions. *N Engl J Med* 2004;350:2196–8.
42. Mauskopf JA, Sullivan SD, Annemans L, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices—budget impact analysis. *Value Health* 2007;10:336–47.

Key Words: costs ■ cost-effectiveness ■ quality of life.